Can a smile reveal your gender?

Piotr Bilinski, Antitza Dantcheva, François Brémond

Can a smile reveal your gender?

Piotr Biliński, Antitza Dantcheva, François Brémond

Inria Méditerranée, 2004 Route des Lucioles - BP 93, 06902 Sophia Antipolis Cedex, France
{Piotr.Bilinski, Antitza.Dantcheva, Francois.Bremond}@inria.fr

***Abstract*—Automated gender estimation has numerous applications including video surveillance, human computer-interaction, anonymous customized advertisement and image retrieval. Most commonly, the underlying algorithms analyze facial appearance for clues of gender. In this work, we propose a novel approach for gender estimation, based on facial behavior in video-sequences capturing smiling subjects. The proposed behavioral approach quantifies gender dimorphism of facial smiling-behavior and is instrumental in cases of (a) omitted appearance-information (*e.g.* low resolution due to poor acquisition), (b) gender spoofing (*e.g.* makeup-based face alteration), as well as can be utilized to (c) improve the performance of appearance-based algorithms, since it provides complementary information. The proposed algorithm extracts spatio-temporal features based on dense trajectories, represented by a set of descriptors encoded by Fisher Vectors. Our results suggest that smile-based features include significant gender-clues. The designed algorithm obtains true gender classification rates of 86.3% for adolescents, significantly outperforming two state-of-the-art appearance-based algorithms (*OpenBR* and *how-old.net*), while for adults we obtain true gender classification rates of 91.01%, which is comparably discriminative to the better of these appearance-based algorithms.**

I. Introduction

Automated gender estimation[1] has drawn high interest for numerous associated applications, ranging from surveillance [36], to human computer-interaction [20], anonymous customized advertisement systems[2], as well as image retrieval systems [2]. At the same time, gender has been a prominent soft-biometric trait [12], [32], [14], which can be employed – in fusion with facial analysis – to improve the matching accuracy of a biometric system [21], in fusion with other soft biometrics for person authentication [10], [11], or it has been employed as a filter for search space reduction [13].

Robust automated gender estimation, currently remains a very challenging task, often due to large intra-class variation [28], and also due to challenges concerning illumination, as well as pose, age and ethnicity of a person. In addition, facial expressions can have a negative effect on the accuracy of automated gender estimation systems. Because of these limitations, the majority of previous works have extracted and studied appearance-based features under the simplifying assumption of constrained pose, expression and illumination, obtaining reasonably good results.

A. Contribution

Deviating from such works, we here propose a method that as opposed to appearance-based methods, analyzes facial behavior, where specifically gender estimation is based on the common facial expression of the *smile*. Given a video-sequence of a smiling subject, the proposed method first extracts dense trajectories from the video-sequence and then represents them by facial spatio-temporal features, which incorporate motion of sampled points in the video-sequence [40], [4], and proceeds to encode these by Fisher Vectors [35]. Based on these, gender of each portrayed subject is classified. Experiments are conducted on the UvA-NEMO dataset [17] containing video-sequences of 400 subjects, where this dataset includes high variation in the subject's age. We compare our results to the performance of two state-of-the-art appearance-based gender estimation algorithms (*OpenBR* [24] and *how-old.net*[3]), and conclude that for adolescents our proposed approach outperforms significantly these state-of-the-art algorithms, while for adults, our algorithm is comparably discriminative.

The premise of our approach is that, generally, male and female smile-dynamics differ in parameters such as intensity and duration [9]. This can be seen in a number of cognitive - psychological studies, which show evidence for gender-dimorphism in the human expression [6], [1], [27], and which observe that females tend to smile more frequently than males in a variety of social contexts [15]. Further, females are more accurate expressers of emotion, when posing deliberately and when observed unobtrusively [5].

Our behavioral based approach is instrumental in cases of (a) omitted appearance - information (*e.g.* due to poor acquisition), (b) gender spoofing (*e.g.* makeup-based gender spoofing [8]), as it is not vulnerable to facial appearance-changes, as well as can serve to (c) improve performance of appearance-based algorithms, since it provides complementary information, and (d) quantify gender-dimorphism of facial smiling-behavior.

B. Structure of paper

This work is organized as follows: Section I-C revisits existing works on gender estimation and expression analysis,

[1] The traditional definition of *sex* refers to the biological characteristics that differentiate men and women, as opposed to *gender*, which is related to the social and cultural distinctions between the sexes. However, very often, "gender" has been used instead of "sex" in biometrics literature and we adopt this annotation.

[2] http://articles.latimes.com/2011/aug/21/business/la-fi-facial-recognition-20110821

[3] http://www.how-old.net/

while Section II describes our proposed method, elaborating on individual steps (dynamic features, encoding, classification). Section III describes the appearance-based state-of-the-art algorithms, Section IV presents the employed dataset, and the subsequent Section V depicts and discusses related experimental results. Finally Section VI concludes the paper.

C. Related work

Existing overview-articles for algorithms related to *gender estimation* include the works of Ng *et al.* [31], Khan *et al.* [22], and Bekios-Calfa *et al.* [2].

Automated Appearance-based Gender Estimation from Face: In gender estimation from face, feature-based approaches extract and analyze a specific set of discriminative facial features (patches) in order to identify the gender of a person. This is a particularly challenging problem, as is implied from the fact that female and male average facial shapes are generally found to be very similar [26].

Another challenge comes in unconstrained settings with different covariates, such as illumination, expressions and ethnicity. While in more constrained settings, face-based gender estimation has been reported to achieve high performance rates, such rates significantly decrease in more realistic and unconstrained settings. An example is the OpenBR gender classification algorithm, that we describe in Section III and utilize later, which was validated on a FERET[4] subset, attaining accuracies of 96.91% and 82.98% for male and female classification, respectively and an overall true classification rate of 90.57% [8], outperforming other algorithms (Neural Network, Support Vector Machines, *etc.*) on the same dataset [29]. In our experiments, given a more unconstrained setting, the algorithm attains true gender classification rates of 52.45% for adolescents and 78.04% for adults.

The majority of gender estimation methods contain two steps preceding face detection, namely *feature extraction* and pattern *classification*. In the context of *feature extraction* Active Appearance Model (AAM) [37], Scale–Invariant Feature Transform (SIFT) [38], Local Binary Patterns (LBP) [41], [29], Semi–Supervised Discriminant Analysis (SDA) [3], as well as combinations of different features [18], [40] have been explored. A number of *classification* methods have been used for gender estimation, and a useful comparative guide of these classification methods can be found in Mäkinen and Raisamo [29].

Body-Dynamics-based Gender Estimation: Dynamics have been used in the context of body-based gender estimation. Related cues include body sway, waist-hip ratio, and shoulder-hip ratio (see [30]); for example, females have a distinct waist-to-hip ratio and swing their hips more, whereas males have broader shoulders and swing their shoulders more.

Despite these recent successes, automated gender recognition from biometric data remains a challenge and is impacted by other soft biometrics, for example, age and ethnicity; gender dimorphism is accentuated only in adults, and varies across different ethnicities.

[4] http://www.nist.gov/itl/iad/ig/colorferet.cfm

II. Proposed Method

Deviating from the above we here propose a novel approach for gender estimation in smile-video-sequences. In what follows, we will describe the basic steps of our proposed method, illustrated in Fig. 1.

We represent the smile-video sequences by local spatio-temporal trajectory-based descriptors. Firstly, we extract dense trajectories in a video sequence. Then, we encode local spatio-temporal volumes around the detected trajectories by motion and appearance descriptors. We then use Fisher vectors, a video representation of motion trajectories, to encode smile-video sequences. Finally, we employ SVM for classification of gender. We proceed to give details on each step.

a) Dense Trajectories:: The dense trajectories approach has been introduced by Wang *et al.* [39]. It extracts local spatio-temporal video trajectories by applying dense sampling of feature points on multiple spatial scales with subsequent tracking of detected feature points using dense optical flow. We extract dense trajectories and proceed to extract local spatio-temporal video volumes around the detected trajectories. We employ dense trajectories for their good coverage of foreground motion and high performance in action recognition.

Trajectory Shape, VCML, HOG, HOF, and MBH descriptors: We extract 5 features aligned with the trajectories to characterize shape (point shifts), appearance (Histogram of Oriented Gradients (HOG) and Video Covariance Matrix Logarithm (VCML)) and motion (Histogram of Optical Flow (HOF) and Motion Boundary Histogram (MBH)). The Trajectory Shape descriptor [39] encodes the shape of a trajectory by a sequence of displacement vectors normalized by the sum of displacement vector magnitudes (see Fig. 2). Four additional descriptors, namely VCML, HOG, HOF, and MBH are computed within a space-time volume around a trajectory. To embed structure information, each local volume is subdivided into a grid of $n_x \times n_y \times n_t$ spatio-temporal cells, where for each cell of the grid, a histogram descriptor is computed. Then, the histograms are normalized with the L_2 norm, and the normalized histograms from cells are concatenated into the final descriptors. The VCML descriptor [4] is based on a covariance matrix representation and it models relationships between different low-level features, such as intensity and gradient. For the HOG and HOF-descriptors, the edge and optical flow orientations are quantized into 8 bins using full orientations, with an additional zero bin for the HOF descriptor. MBH divides the optical flow field $I_w = (I_x, I_y)$ into x and y components, spatial derivatives are computed separately for the horizontal and vertical components of the optical flow, and orientation information is quantized into histograms, similarly to the HOG descriptor. The MBH descriptor encodes the relative motion between pixels. Constant motion information is suppressed and only information about changes in the flow field (*i.e.*, motion boundaries) is kept.

In the following, we utilize the parameters of spatial size of the volume 32×32, and n_x, n_y = 3, and $n_t = 2$, as in [40], [4].

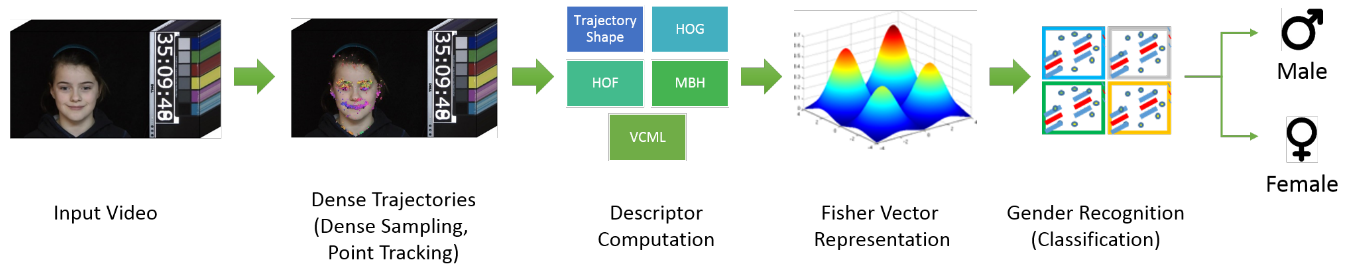
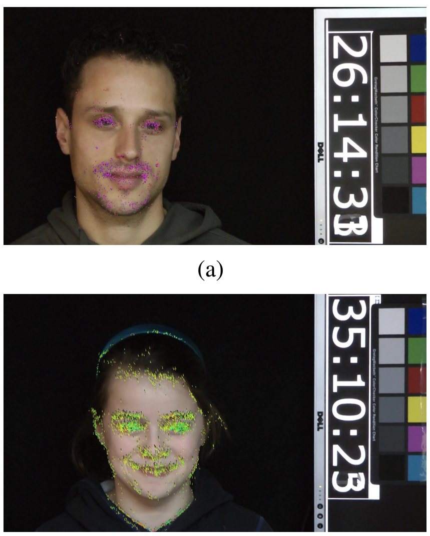


Fig. 1: Proposed framework for appearance and dynamic gender estimation. Local spatio-temporal features for dense trajectories are computed and the corresponding descriptors are encoded with Fisher Vectors. Late fusion is applied to get final smile representation and SVM is utilized for gender recognition.



(a)

(b)

Fig. 2: Dense trajectories extracted during a smile expressed by a male and female.

b) Smile Representation with Fisher Vector encoding:: We apply the Fisher Vector encoding to represent smile-video sequences using the extracted motion trajectories and their corresponding descriptors.

Fisher Vector encoding has been introduced by Perronnin and Dance [33] and it has been improved by Perronnin *et al.* in [35]. It is a video (and image) descriptor obtained by pooling local features into a global representation, and it can be seen as an extension of the widespread bag-of-features approach. It describes local features by their deviation from a "universal" generative Gaussian Mixture Model (GMM). It has shown to achieve excellent results as a global descriptor both for image classification [35] and for image retrieval [34], outperforming standard bag-of-features approach.

We apply the Fisher Vector encoding with first and second order statistics of local features [35] to represent videos using local descriptors. We calculate a separate video representation for each descriptor independently (*i.e.*, Trajectory Shape, HOG, HOF, MBH and VCML). Then, for each video representation, we apply the power normalization and the L2 norm, as in [35]. Finally, to combine different descriptor types, we concatenate their normalized Fisher Vectors (*i.e.*, per video: we concatenate the Fisher Vector representation from HOG with Fisher Vector representation from HOF, *etc.*). The final video representation is of size $2DK$, where D is the sum of dimensions of descriptors and K is the number of Gaussians.

c) Gender Classification with Support Vector Machines:: For gender classification, we use linear Support Vector Machines (SVM) [7]. The motivations behind using this classifier are twofold. First, it has shown to provide very good and promising results with high-dimensional data such as Fisher vectors, as typically if the number of features is large, there is no need to map data to a higher dimensional space [19]. Second, it has shown to be very efficient both in training and prediction steps.

d) Implementation Details:: For the Fisher Vector, we apply the Principal Component Analysis (PCA) before the encoding, and we reduce a descriptor dimensionality by a factor of two, as suggested by Perronnin *et al.* [35]. To estimate the GMM parameters, we randomly sample a subset of $100,000$ features from the training set. We consider 3 codebook-sizes (64, 128 and 256 Gaussians) and we set the codebook size using cross-validation. To increase precision, we initialize the GMM ten times and we keep the codebook with the lowest error. Towards evaluating the performance of our proposed algorithm, we use five-folds cross-validation, where we try to balance the number of male to female video instances. Per split, we calculate mean person accuracy (*i.e.*, mean accuracy from various video instances of a person, if applicable) and we report the average value from all people belonging to this split. Final accuracy (True Gender Classification Rate) is obtained by averaging the accuracy from all the splits.

Having described our proposed dynamics-based method, we proceed to specify two baseline appearance-based algorithms that were used for comparison of the performance of our algorithm.

III. Baseline Algorithms

We compare the proposed algorithm with following two appearance-based state-of-the-art algorithms.

OpenBR [24] is a publicly available open source software for biometric recognition and evaluation. We utilize the gender estimation algorithm, based on the work of Klare *et al.* [23]. Specifically, a face image is represented by extracting histograms of LBP and SIFT-features computed on a dense grid of patches. Subsequently, the histograms from each patch are projected onto a subspace generated using PCA in order to obtain a feature vector. SVM is used for classification.

how-old.net is a website (http://how-old.net/) launched by Microsoft for online age and gender recognition. Images can be uploaded and as an output age and gender labels are provided. The underlying algorithm and training dataset are not publicly disclosed.

Since the video-sequences of the UvA-NEMO dataset start with the neutral expression of the portrayed subject, the first frame is utilized to extract appearance features, as expressions have shown to influence gender recognition performance of appearance-based methods [37].

We proceed to elaborate on the dataset employed in the gender estimation experiments.

IV. Dataset

The UvA-NEMO Smile Dataset[5], introduced by Dibeklioğlu *et al.* [17], consists of 1-2 video sequences of 400 subjects (185 females, 215 male). In total, UvA-NEMO contains 597 video sequences (294 female instances, 303 male instances). To elicit spontaneous smiles, each subject was displayed a short funny video segment. Each video starts and ends with neutral or a near-neutral expression of the subject (see Fig. 3). The pose of the subjects is frontal, camera-to-subject and illumination are reasonably constant across subjects. The resolution of the videos is 1920×1080 pixels at a framerate of 50 frames per second. The right side of the images includes time-indication, that we have cropped out in our experiments.

The UvA-NEMO dataset consists of images of subjects in the age-range of 8 to 76 years. The ability of dynamics to predict age, and thus the impact of age on a small set of facial dynamics has been previously assessed in the work of Dibeklioğlu *et al.* [16], where results suggest that facial-dynamics change significantly for adolescents and adults. Consequently we present our results based on two age-categories.

We note that the ethnicity of subjects in the UvA-NEMO dataset is predominantly Caucasian, hence the current study does not reflect on covariates such as ethnicity, as well as social and cultural background.

V. Experiments

Towards evaluating the performance of the proposed gender estimation algorithm, we employ a 5-fold cross-validation scheme. Here, the UvA-NEMO dataset is divided into 5 folds, 4 folds are used for training, and the remaining fold is used for testing it. This is repeated 5 times and reported results are the average thereof. Note that the subjects in the test set are not present in the training set.

In Table I we report the gender estimation accuracy of the proposed algorithm and the two baseline algorithms. We observe that the baseline *appearance based gender algorithms* perform relatively well for the age category > 20 years and rather poorly in the age category ≤ 20 years. This can be due to no facial sexual dimorphism for toddlers and adolescents [25] (see Figure 4). Interestingly, our proposed algorithm (True Gender Classification Rate $TGCR = 86.3\%$) outperforms the appearance based algorithms (OpenBR: $TGCR = 52.35\%$ and *how-old.net*: $TGCR = 55.55\%$) significantly in the age category ≤ 20 years. In the age category > 20, *how-old.net* performs best ($TGCR = 92\%$), with comparable results obtained by the proposed algorithm ($TGCR = 91.01\%$). The results suggest that facial smile-dynamics carry substantial cues related to gender of the subject. To undermine this statement, we have computed classification results related to the merely motion-based descriptors. Jointly the trajectory shape, HOF and MBH obtain a $TGCR = 77.7\%$ for adolescents and $TGCR = 80.11\%$ for adults.

Tab. I: True gender classification rates of the proposed method and two state-of-the-art gender estimation appearance based algorithms. Age given in years.

Age	≤ 20	> 20
Subject amount	148	209
OpenBR [24]	52.35%	75.58%
how-old.net	55.55%	$\mathbf{92}\%$
Motion-based descriptors	77.7%	80.11%
Proposed Method	$\mathbf{86.3}\%$	$\mathbf{91.01}\%$

The classification rate for each trial in the 5-fold cross-validation experiment is reported for subjects ≤ 20 years old in Table II and for subjects > 20 years old in Table III. We observe that our algorithm consistently outperforms the appearance-based algorithms for all trials in the age-group ≤ 20, with a high peak in the trial 4, where interestingly the appearance-based algorithms have a low peak. Correspondingly, the low peaks of our algorithm (trial 1 and 3) are high peaks of the appearance-based algorithms, respectively, which indicates the complementary nature of appearance and dynamic-based algorithms. This cannot be observed in the age-group > 20 though. We observe that our algorithm outperforms OpenBR and performs comparably to *how-old.net*.

VI. Conclusions

In this work we proposed a novel gender estimation approach, which analyzes and classifies facial smiling-behavior in video-sequences. Such an approach has the advantage of being complementary to appearance-based approaches, and thus being robust to gender spoofing. The proposed algorithm utilizes dense trajectories represented by spatio-temporal facial features and Fisher Vector encoding. Our results suggest that for subjects of less or equal to 20 years old, our behavioral-features approach significantly outperforms the existing state-of-the-art appearance-based algorithms, while for subjects

[5]http://www.uva-nemo.org

(a) (b) (c) (d) (e) (f)

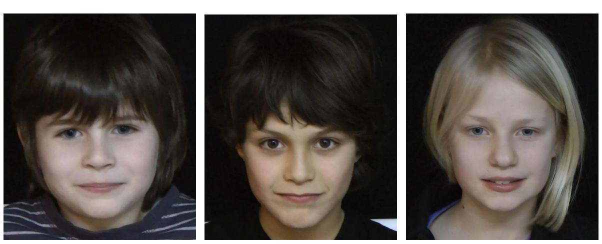
Fig. 3: Example male and female subjects from the UvA-NEMO dataset expressing spontaneous smiles. (a), (d) First frames, (b), (e) intermediate frames, and (c), (f) last frames of example video sequences.

Tab. II: True gender classification rates of proposed method and two state-of-the-art gender estimation appearance based algorithms. Subjects ≤ 20 years old. The numbers in parentheses indicate the number of male and female video-sequences in each trial.

Trial	Train	Test	OpenBR	*how-old.net*	Proposed Method
1	199 (100/99)	48(24/24)	50%	56.25%	82.76%
2	197 (99/98)	50(25/25)	52%	56%	87.93%
3	201 (101/100)	46(23/23)	56.86%	54.9%	82.76%
4	196 (98/98)	51(26/25)	49.02%	52.94%	91.38%
5	195 (98/97)	52(26/26)	53.85%	57.69%	86.67%
Average			52.35%	55.55%	86.3%

Tab. III: True gender classification rates of proposed method and two state-of-the-art gender estimation appearance based algorithms. Subjects > 20 years old. The numbers in parentheses indicate the number of male and female video-sequences in each trial.

Trial	Train	Test	OpenBR	*how-old.net*	Proposed Method
1	278 (142/136)	72(37/35)	80.56%	90.28%	95.24%
2	283 (145/138)	67(34/33)	85.07%	89.55%	86.9%
3	278 (142/136)	72(37/35)	77.8%	93.06%	91.67%
4	285 (145/140)	65(33/32)	72.31%	93.85%	92.86%
5	276 (141/135)	74(36/38)	62.16	93.24%	88.37%
Average			75.58%	92%	91.01%



(a) (b) (c)

Fig. 4: Example subjects of the age-group ≤ 20 years, where facial appearance is not a clear indicator for the gender, from human and machine point of views.

above 20 years old, our approach is comparatively discriminative to the state-of-the-art algorithms. Future work will seek to estimate soft biometrics analyzing additional facial expressions.

Acknowledgements: A. Dantcheva was supported as an ERCIM Alain Bensoussan Fellow at INRIA funded by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement #246016 and as a Postdoctoral Fellow at INRIA funded by the French Government (National Research Agency, ANR) through the "Investments for the Future" Program reference #ANR-11-LABX-0031-01.

References

[1] R. B. Adams, U. Hess, and R. E. Kleck. The intersection of gender-related facial appearance and facial displays of emotion. *Emotion Review*, 7(1):5–13, 2015.

[2] J. Bekios-Calfa, J. M. Buenaposada, and L. Baumela. Revisiting linear discriminant techniques in gender recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 33(4):858–864, 2011.

[3] J. Bekios-Calfa, J. M. Buenaposada, and L. Baumela. Robust gender recognition by exploiting facial attributes dependencies. *Pattern Recognition Letters*, 2013.

[4] P. Bilinski and F. Bremond. Video covariance matrix logarithm for human action recognition in videos. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 2015.

[5] L. R. Brody and J. A. Hall. Gender and emotion in context. *Handbook of emotions*, 3:395–408, 2008.
[6] E. Cashdan. Smiles, speech, and body posture: How women and men display sociometric status and power. *Journal of Nonverbal Behavior*, 22(4):209–228, 1998.
[7] C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27:1–27:27, 2011.
[8] C. Chen, A. Dantcheva, and A. Ross. Impact of facial cosmetics on automatic gender and age estimation algorithms. In *International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP)*, 2014.
[9] A. Dantcheva and F. Brémond. Gender estimation based on smile-dynamics. *IEEE Transactions on Information Forensics and Security (TIFS)*, 2016.
[10] A. Dantcheva, J.-L. Dugelay, and P. Elia. Person recognition using a bag of facial soft biometrics (BoFSB). In *IEEE International Workshop on Multimedia Signal Processing (MMSP)*, 2010.
[11] A. Dantcheva, J.-L. Dugelay, and P. Elia. Soft biometric systems: reliability and asymptotic bounds. In *IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, 2010.
[12] A. Dantcheva, P. Elia, and A. Ross. What else does your biometric data reveal? a survey on soft biometrics. *IEEE Transactions on Information Forensics and Security (TIFS)*, 11(3):441–467, 2016.
[13] A. Dantcheva, A. Singh, P. Elia, and J.-L. Dugelay. Search pruning in video surveillance systems: Efficiency-reliability tradeoff. In *IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, 2011.
[14] A. Dantcheva, C. Velardo, A. D'Angelo, and J.-L. Dugelay. Bag of soft biometrics for person identification. New trends and challenges. *Multimedia Tools and Applications (MTAS)*, 51:739–777, 2011.
[15] F. M. Deutsch, D. LeBaron, and M. M. Fryer. What is in a smile? *Psychology of Women Quarterly*, 11(3):341–352, 1987.
[16] H. Dibeklioğlu, T. Gevers, A. A. Salah, and R. Valenti. A smile can reveal your age: Enabling facial dynamics in age estimation. In *ACM International Conference on Multimedia*, pages 209–218, 2012.
[17] H. Dibeklioğlu, A. A. Salah, and T. Gevers. Are you really smiling at me? spontaneous versus posed enjoyment smiles. In *European Conference on Computer Vision (ECCV)*, pages 525–538. Springer, 2012.
[18] G. Guo, C. Dyer, Y. Fu, and T. Huang. Is gender recognition affected by age? In *International Conference on Computer Vision Workshops (ICCV Workshops)*, pages 2032–2039, 2009.
[19] C.-W. Hsu, C.-C. Chang, and C.-J. Lin. A practical guide to support vector classification. Technical report, Department of Computer Science, National Taiwan University, 2003.
[20] J. A. Jacko. *Human computer interaction handbook: Fundamentals, evolving technologies, and emerging applications*. CRC press, 2012.
[21] A. K. Jain, S. C. Dass, and K. Nandakumar. Can soft biometric traits assist user recognition? In *SPIE Defense and Security Symposium*, volume 5404, pages 561–572, 2004.
[22] S. A. Khan, M. Nazir, S. Akram, and N. Riaz. Gender classification using image processing techniques: a survey. In *International Multi Topic Conference (INMIC)*, 2011.
[23] B. F Klare, M. J. Burge, J. C. Klontz, R. W. Vorder Bruegge, and A. K. Jain. Face recognition performance: Role of demographic information. *IEEE Transactions on Information Forensics and Security (TIFS)*, 7(6):1789–1801, 2012.
[24] J. C. Klontz, B. F. Klare, S. Klum, A. K. Jain, and M. J. Burge. Open source biometric recognition. In *IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–8, 2013.
[25] J. Koudelova, J. Bruzek, V. Caganova, V. Krajicek, and J. Veleminska. Development of facial sexual dimorphism in children aged between 12 and 15 years: a three-dimensional longitudinal study. *Orthodontics & craniofacial research*, 18(3):175–184, 2015.
[26] J. H. Langlois and L. A. Roggman. Attractive faces are only average. *Psychological science*, 1(2):115–121, 1990.
[27] M.-F. Liébart, C. Fouque-Deruelle, A. Santini, F. Dillier, V. Monnet-Corti, J.-M. Glise, and A. Borghetti. Smile line and periodontium visibility. *Perio*, 1(1):17–25, 2004.
[28] S. R. Loth and M.Y. Iscan. *Sex determination, Encyclopedia of Forensic Sciences*, volume 1. Academic Press, San Diego, 2000.
[29] E. Mäkinen and R. Raisamo. Evaluation of gender classification methods with automatically detected and aligned faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 30(3):541–547, 2008.
[30] G. Mather and L. Murdoch. Gender discrimination in biological motion displays based on dynamic cues. In *Biological Sciences B*, pages 273–279, 1994.
[31] C. B. Ng, Y. H. Tay, and B.-M. Goi. Vision-based human gender recognition: A survey. *PRICAI 2012: Trends in Artificial Intelligence. Lecture Notes in Computer Science*, 7458:335–346, 2012.
[32] M. S. Nixon, P. L. Correia, K. Nasrollahi, T. B. Moeslund, A. Hadid, and M. Tistarelli. On soft biometrics. *Pattern Recognition Letters*, 68, Part 2:218 – 230, 2015.
[33] F. Perronnin and C. Dance. Fisher kernels on visual vocabularies for image categorization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8. IEEE, 2007.
[34] F. Perronnin, Y. Liu, J. Sánchez, and H. Poirier. Large-scale image retrieval with compressed fisher vectors. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3384–3391. IEEE, 2010.
[35] F. Perronnin, J. Sánchez, and T. Mensink. Improving the Fisher Kernel for large-scale image classification. In *European Conference on Computer Vision (ECCV)*, pages 143–156. 2010.
[36] D. Reid, S. Samangooei, C. Chen, M. Nixon, and A. Ross. Soft biometrics for surveillance: An overview. In *Handbook of Statistics*, volume 31, 2013.
[37] Y. Saatci and C. Town. Cascaded classification of gender and facial expression using Active Appearance Models. In *IEEE International Conference on Automatic Face and Gesture Recognition (FGR)*, pages 393–400, 2006.
[38] M. Toews and T. Arbel. Detection, localization, and sex classification of faces from arbitrary viewpoints and under occlusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 31(9):1567–1581, 2009.
[39] H. Wang and C. Schmid. Action recognition with improved trajectories. In *IEEE International Conference on Computer Vision (ICCV)*, pages 3551–3558, 2013.
[40] J. Wang, J. Li, W. Yau, and E. Sung. Boosting dense sift descriptors and shape contexts of face images for gender recognition. In *Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 96–102, 2010.
[41] Z. Yang and H. Ai. Demographic classification with local binary pattern. In *IAPR International Conference on Biometrics (ICB)*, 2007.